

ARTYKUŁY POGLĄDOWE (REVIEW PAPERS)

Problemy w upowszechnieniu analiz dużych skupisk informacji w ochronie zdrowia

(Problems in dissemination of analyses of large concentrations of information in health care)

S Kasza^{1,A,D}, A Romaszewski^{1,E}, Z Kopański^{1,D,F}, W Uracz^{2,B,E}, F Furmanik^{2,C}, S Dyl^{2,B}, J Tabak^{2,B}

1. Wydziału Nauk o Zdrowiu Collegium Medicum Uniwersytet Jagielloński
2. Collegium Masoviense – Wyższa Szkoła Nauk o Zdrowiu

Abstract - Currently, the number of data produced and collected by computer science is increased many times in broadly understood health care. Most of the information available today was still impossible to obtain a decade ago, and the data growth will continue. Today, organizational systems around the world are constantly evolving. This is also the case with the health protection system, which as a relatively young structure develops very dynamically. Factors such as technological progress, social or political changes have a continuous impact on the structure of the mechanisms of operation and the assumption of the system. Contemporary tendencies to digitize administration and data digitization also do not bypass this sector. Worldwide, work has been carried out for a long time in order to fully informatisation at the level of individual medical entities as well as entire countries. These issues devote more attention to these issues. Against this background, they discuss the elements of implementation and functioning of the Big Data system in health care. They characterize the role of cloud computing and its usefulness in the work of medical facilities.

Key words - health protection, Big Data system, cloud computing.

Streszczenie - Obecnie liczba produkowanych oraz gromadzonych danych przez informatykę zwiększa się wielokrotnie w szeroko rozumianej ochronie zdrowia. Większość dostępnych dzisiaj informacji była jeszcze niemożliwa do uzyskania dekadę temu, a przyrost danych będzie w dalszym ciągu postępował. Współcześnie systemy organizacyjne na całym świecie podlegają ciągłej ewolucji. Dzieje się tak też w przypadku systemu ochrony zdrowia, który jako relatywnie młoda struktura rozwija się bardzo dynamicznie. Czynniki takie jak postęp technologiczny, zmiany społeczne czy polityczne wywierają ciągły wpływ na strukturę mechanizmów działania i założenia systemu. Współczesne tendencje do cyfryzacji administracji i digitalizacji danych nie omijają także tego sektora. Na całym świecie od dłuższego czasu są prowadzone prace mające na celu dokonać pełnej informatyzacji na poziomie indywidualnych podmiotów medycznych jak i całych państw. Tym zagadnieniom autorzy poświęcają więcej uwagi. Na tym tle omawiają elementy wdrożenia i funkcjonowania systemu Big Data w ochronie zdrowia. Charakteryzują rolę chmury obliczeniowej i jej przydatność w pracy placówek medycznych.

Słowa kluczowe - ochrona zdrowia, systemu Big Data, chmura obliczeniowa.

Wkład poszczególnych autorów w powstanie pracy - A-Koncepcja i projekt badania, B-Gromadzenie i/lub zestawianie danych, C-Analiza i interpretacja danych, D-Napisanie artykułu, E-Krytyczne zrecenzowanie artykułu, F-Ostateczne zatwierdzenie artykułu

Adres do korespondencji — Prof. dr Zbigniew Kopański, Wydziału Nauk o Zdrowiu Collegium Medicum Uniwersytet Jagielloński, Kraków, ul. Piotra Michałowskiego 12, PL-31-126 Kraków, e-mail: zkopanski@o2.pl

Zaakceptowano do druku: 29.08.2018.

WSTĘP

Trudno jednoznacznie wskazać początek zjawiska przechowywania i przetwarzania na większą skalę danych medycznych i informacji związanych z pacjentem. Prawdopodobnie pierwsze tego typu

próby miały miejsce w 1883 roku, kiedy wprowadzono w Niemczech pierwszy publiczny model ubezpieczenia zdrowotnego Bismarcka. Wraz ze wprowadzeniem systemu stało się niezbędnym ewidencjonowanie osób uprawnionych do otrzymywania świadczeń. Na początku rozwoju systemów opieki

zdrowotnej nie przykładano jednak uwagi do prywatności i poufności przechowywania danych. Sytuacja uległa zmianie na przestrzeni lat z uwagi na szeroki dostęp do publicznych usług finansowanych często z budżetu państwa. Nie mniej idea ochrony danych osobowych na dobre zadomowiła się w umysłach ludzi dopiero pod koniec XX wieku wraz rozwojem społeczeństwa informacyjnego, gdzie wiedza oraz informacja stanowią równie istotną wartość, a niejednokrotnie wyższą niż pieniądze. [1,2]

GROMADZENIE DANYCH W OCHRONIE ZDROWIA

Obecnie liczba produkowanych oraz gromadzonych danych przez informatykę zwiększa się wielokrotnie w szeroko rozumianej ochronie zdrowia. Większość dostępnych dzisiaj informacji była jeszcze niemożliwa do uzyskania dekadę temu, a przyrost danych będzie w dalszym ciągu postępował. Współcześnie systemy organizacyjne na całym świecie podlegają ciągłej ewolucji. Dzieje się tak też w przypadku systemu ochrony zdrowia, który jako relatywnie młoda struktura rozwija się bardzo dynamicznie. Czynniki takie jak postęp technologiczny, zmiany społeczne czy polityczne wywierają ciągły wpływ na strukturę mechanizmów działania i założenia systemu. Współczesne tendencje do cyfryzacji administracji i digitalizacji danych nie omijają także tego sektora. Na całym świecie od dłuższego czasu są prowadzone prace mające na celu dokonać pełnej informatyzacji na poziomie indywidualnych podmiotów medycznych jak i całych państw. Dotyczy to także sposobu przetwarzania danych. Pojedyncze placówki ochrony zdrowia przy odpowiednim nakładzie środków już teraz są w stanie opracowywać gromadzone wielkie zasoby danych. Zabieg ten pozwala zwiększenie konkurencyjności placówki na rynku poprzez przewidywanie zdarzeń i efektywne wykorzystanie zasobów. Natomiast na poziomie państwowym, pierwsze sukcesy wdrożeniowe krajów zachodnich takich jak Dania [3], Szwecja czy Australia, pokazują, że udana implementacja nowatorskich technologii, skutkuje ulepszeniem całego systemu zdrowotnego. W szczególności bierze się pod uwagę wzrost wydajności systemu oraz - zgodnie z teorią - oszczędności, ponieważ koszty administracji maleją. Elektroniczne zbieranie, przetwarzanie i wykorzystywanie danych ma też szerokie zastosowanie w analizach stanu zdrowotnego populacji i demografii, stronie państwowej pozwala lepiej kreować politykę

zdrowotną, a także zapewniać rozwiązania według potrzeb i reagować na zmieniającą się dynamicznie sytuację w kraju. Wszystkie te zadania niezależnie od szczebla, na którym się odbywają, wymagają jednak szerokiej bazy danych, która będzie gromadziła informacje o uczestnikach placówki lub systemu zdrowotnego. Tego typu sytuacja rodzi wiele pytań o bezpieczeństwo i sposób użycia zgromadzonych informacji. Powstają też problemy związane z pozyskiwaniem danych medycznych, a także ich nadzoru i kontroli przez samych obywateli. Mnogość administratorów danych, a także możliwość ich potencjalnego wykorzystania wymusza wprowadzenie uniwersalnych praw dla jednostek udzielających pozwolenia do zbierania, przetwarzania i wykorzystywania danych. Pełna kontrola obywatela nad prawidłowością przetwarzanych danych zapewnia najlepsze zabezpieczenie systemu, a także osób fizycznych przed nieuprawnionym lub błędnym wykorzystaniem zgromadzonych informacji.

Analiza wielkich zbiorów danych pozyskiwanych z różnych źródeł, daje nieograniczone możliwości osiągnięcia nowej wiedzy, istotnej w podejmowaniu kluczowych decyzji. W długim okresie czasu nowe informacje zdobyte dzięki nowoczesnym metodom analizy danych, przełożą się pośrednio na lepsze funkcjonowanie i opiekę w ochronie zdrowia, a także redukcję kosztów systemu. Umożliwią wgląd w powody i skutki chorób, lepszą precyzję w doborze leków oraz wzmocnią przewidywanie oraz prewencję chorób. Skorzysta zatem pacjent, który otrzyma lepszą, bardziej sprofilowaną opiekę zdrowotną. Big Data mogą udoskonalić nasze zrozumienie na temat zachowań zdrowotnych takich jak palenie tytoniu, picie alkoholu i wielu innych oraz przyspieszyć dyfuzję uzyskanych innowacji. [1,2 4]

WDROŻENIU I FUNKCJONOWANIU SYSTEMU BIG DATA

Pomimo wielu technicznych trudności we wdrożeniu i funkcjonowaniu systemu analiz Big Data, spora liczba przedsiębiorstw z różnych sektorów gospodarki pierwsze poważne kroki ma już za sobą. Z powodzeniem wykorzystują pierwsze analizy do określania wszelakich wzorców, które następnie z sukcesem wykorzystują w swoich działaniach rynkowych. Obecnie dzieje się to bez względu na to czy źródło informacji jest publiczne czy prywatne. Przedsiębiorstwa dzięki swoim działaniom nie tylko zwiększają swoje zyski, ale też zwiększają satys-

fakcje konsumenta. Niestety tradycyjnie system ochrony zdrowia pozostaje nieznacznie w tyle w temacie używania nowych technologii w tym Big Data. Po pierwsze po części wynika to z faktu oporności na zmiany. Świadczeniodawcy przywykli do podejmowania decyzji niezależnie, wykorzystując kliniczny osąd zamiast polegać na protokołach bazujących na statystycznej analizie danych. Kolejne przeszkody są bardziej przyziemne. Świadczeniodawcy często unikają inwestycji, które są niepewne i nie zrozumiałe, a spodziewane zyski z nakładów są rozmyte i niejednoznaczne. Instytucje zdrowotne wolą pozostać przy sprawdzonych systemach, po mimo dużych ograniczeń w przetwarzaniu gromadzonych danych. Kolejnym powodem ograniczającym rozwój analiz wielkich baz danych jest po prostu brak nowoczesnych oraz łatwych w obsłudze narzędzi. Pojawiają się pierwsze programy tego typu, ale ich cena w stosunku do możliwości jest wciąż zbyt wysoka, aby małe podmioty z tego sektora mogły sobie na nią pozwolić. Istnieje też problem kompetencji. W tym momencie nie ma na rynku wielu specjalistów, którzy mieliby dostateczną wiedzę lub umiejętności do pracy ze specyficznymi danymi medycznymi. Wykształcenie w zakresie analizy wielkich baz danych w ochronie zdrowia może stanowić problem, ponieważ jest to dziedzina należąca do grupy interdyscyplinarnych, która wymaga poznania wielu różnych zagadnień z dziedziny informatyki, statystyki oraz medycyny. Dużą przeszkodą dla rozwoju zaawansowanej medycyny informatycznej jest również kwestia osobliwego środowiska medycznego. Mnogość podmiotów na rynku o różnych zadaniach oraz obowiązkach wymaga skomplikowanego i skoordynowanego sposobu przesyłu i wymiany danych. W głównej mierze na drodze do dynamicznego rozwoju stoi kwestia prywatności. Dane medyczne oraz dane osobowe zawarte w systemie są zazwyczaj chronione w sposób szczególny. Mówi się często o danych wrażliwych, których udostępnianie, gromadzenie oraz przetwarzanie są obłożone rozmaitymi obostrzeniami prawnymi zarówno na szczeblu narodowym jak i na szczeblu ponadpaństwowym, porozumień międzynarodowych. Dotyczy to także wymiany informacji wewnątrz pojedynczej placówki szpitalnej, firmy farmaceutycznej czy innego podmiotu zaangażowanego w ochronę zdrowia. Brak wypracowanych do tej pory ściśle określonych procedur rzutuje na funkcjonowanie całego systemu i perspektywach rozwoju. Wejście w nową epokę tzw. erę informacji gwarantuje jednak coraz szybszy i bardziej

dynamiczny rozwój także tej branży. Gwałtowna seria trendów w dziedzinie informatyzacji otaczającego nas świata, pozwala jednak przypuszczać, że również branża medyczna dojrzała do zmian w organizacji. Według autorów raportu „Big Data Revolution”, branża w chwili obecnej osiągnęła punkt krytyczny i jest gotowa do zmian, także do implementacji Big Data. Gotowość do wprowadzenia nowych technologii wiąże się z wysokimi kosztami funkcjonowania obecnego systemu, klimatem ekonomicznym i powszechnymi zmianami rynkowymi [5-8].

Tabela 1. Zestawienie korzyści i przeszkód wdrożenia analiz Big Data w ochronie zdrowia [Opracowanie własne na podstawie 9]

Problemy w upowszechnieniu analiz dużych skupisk informacji w ochronie zdrowia	
Korzyści	Przeszkody
<ul style="list-style-type: none"> • Nowy rodzaj wiedzy; • Sprawniejsze funkcjonowanie i opieka w ochronie zdrowia; • Redukcja kosztów systemu; • Profilowane leczenie; • Nowe metody leczenia i diagnostyki; Lepsza profilaktyka w populacji; • Lepsze zrozumienie zachowań zdrowotnych; 	<ul style="list-style-type: none"> • Oporność na zmiany; • Niejasne korzyści z inwestycji; • Brak prostych narzędzi; • Zbyt wysoka cena dostępnego oprogramowania; • Niedobór specjalistów z zakresu przetwarzania danych w ochronie zdrowia; • Niejednorodna struktura podmiotów korzystających z danych; • Nieprecyzyjne i skomplikowane prawo krajowe oraz międzynarodowe bądź jego brak;

CHMURA OBLICZENIOWA

Chmura obliczeniowa jako przyszłość w ochronie zdrowia Chmura obliczeniowa (ang. cloud computing) „*jest usługą polegającą na zdalnym udostępnieniu mocy obliczeniowej urządzeń IT oferowaną przez zewnętrzne podmioty, dostępną na żądanie w dowolnej chwili oraz skalującą się w miarę zapotrzebowania*.” [10] Chociaż chmury obliczeniowe nie są nową koncepcją i są powszechnie wykorzystywane w dzisiejszych rozwiązaniach internetowych w różnych modelach to w ochronie zdrowia stanowią nową ideę. Chmurę obliczeniową można zdefiniować pod postacią 5 charakterystycznych cech takich jak [11-15]:

- mierzalność usługi;

- samoobsługa na żądanie;
- nieograniczony dostęp do sieci;
- występowanie puli zasobów;
- elastyczność i szybkość usługi.

Najprostszym przykładem przedstawiającym jeden z modeli chmury jest poczta elektroniczna [15].

Akumulowanie oraz przetwarzanie znaczącej liczby faktów niejednokrotnie wymaga ponadprzeciętnej mocy obliczeniowej liczonej w setkach lub tysiącach serwerów. Jedynie nieliczne przedsiębiorstwa mogą zainwestować w infrastrukturę zdolną udźwignąć tego rodzaju zadania. Kosztowne oraz skomplikowane rozwiązania mogą zostać zastąpione przez usługi tzw. przetwarzania danych w chmurze. Jedną z licznych zalet tego rozwiązania jest skalowalność. Wraz ze wzrostem posiadanych informacji nie dochodzi do rozwoju infrastruktury po stronie usługobiorcy. Komercyjne z tego rodzaju propozycją występują firmy takie jak Google, Amazon czy Microsoft, które stworzyły rozwiązania na własne potrzeby analizy dużych skupisk danych, ale w związku z rosnącym popytem, postanowiły zaoferować je szerokiej publiczności.

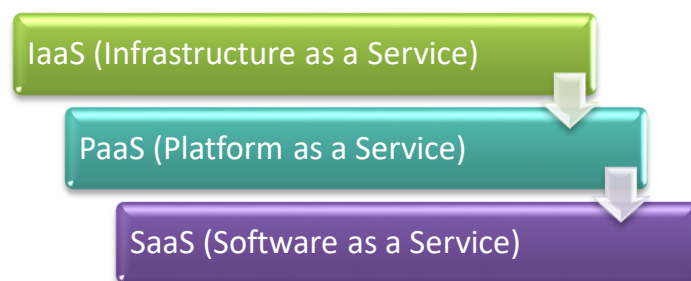
W kontekście licznych problemów związanych z gromadzeniem i przetwarzaniem dużych ilości danych w ochronie zdrowia, powstają rozwiązania technologiczne, które mają umożliwić oraz ułatwić pracę z danymi. Na tle obecnych rozwiązań na świecie, najbardziej efektywnym oraz ekonomicznym modelem operacji na danych jest schemat zmodyfikowanej dla potrzeb publicznych chmury obliczeniowej, idee stanowiącą wielofunkcyjny magazyn wykorzystywany do pracy na powierzonych danych.

Rozpowszechnienie tego rodzaju usługi dla podmiotów funkcjonujących w ochronie zdrowia jest korzystne z przyczyn praktycznych, a także ekonomicznych. Placówki medyczne, w których odbywa się proces przetwarzania danych o stanie zdrowia mając na uwadze zmieniające się w Polsce, a także Unii Europejskiej przepisy, nie będą musiały tworzyć odrębnej infrastruktury i rozwiązań dla placówki. Zamiast tego będą mogły skorzystać z firm zewnętrznych, które przez specjalizację swoich usług, są w stanie zaoferować produkty solidne, dostosowane do norm prawa i po konkurencyjnej cenie. Wspomniane rozwiązanie jest istotne w związku z nowymi wymogami stawianymi przez system informacji w ochronie zdrowia. Nowe narzędzia służące do identyfikacji i uwierzytelniania dokumentów będące w polskim ustawodawstwie

jeszcze w trakcie koncepcji, czyli elektroniczna karta ubezpieczenia zdrowotnego, a także karta specjalisty medycznego mogą w niedługim czasie stanowić kluczowe narzędzia nowego systemu generującego duże ilości danych. Celem nowych narzędzi będzie min. podpisanie elektronicznej dokumentacji medycznej podpisem elektronicznym oraz informowanie modułu Systemu Informacji Medycznej o dokonanych działaniach. Zadaniem min. chmury obliczeniowej będzie koordynowanie i przetwarzanie, gromadzenie i archiwizacja informacji pochodzących z wielu źródeł. Niestety obecnie ramy prawne dla nowoczesnych rozwiązań nie zostały jeszcze dokładnie określone, jednak istnieją już pierwsze zarysy nowych zasad dla systemu informacji medycznej w Polsce, które zostaną poniżej pokrótce opisane [11-18].

ZASTOSOWANIE CHMURY OBLICZENIOWEJ

Zastosowanie Szerokie możliwości zastosowania chmur obliczeniowych w każdym sektorze gospodarki w tym na rynku ochrony zdrowia oznacza, że rozwiązanie posiada wiele rodzajów usług, które są mogą być wykorzystywane w zależności od rozmaitych potrzeb. Na chwilę obecną można wyróżnić kilka rozwiązań chmurowych, co przedstawiono na rycinie 1.



Rycina 1. Przykładowe rozwiązania chmurowe [18]

IaaS (Infrastructure as a Service) to rozwiązanie najprostsze. Polega na udostępnianiu przez zewnętrznego dostawcę samej infrastruktury sprzętowej: przestrzeni dyskowej, mocy obliczeniowej sieci.

PaaS (Platform as a Service) jest rozwiązaniem trochę szerszym, obejmującym też udostępnianie oprogramowania pośredniczącego oraz narzędzi do rozwijania własnych aplikacji (przede wszystkim będzie to Hadoop i MapReduce).

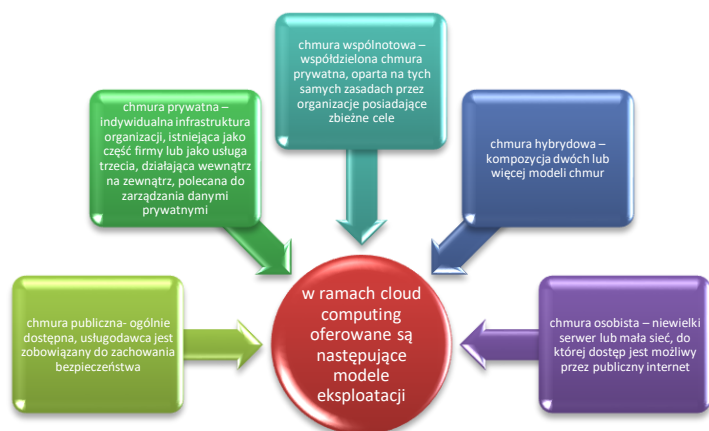
SaaS (Software as a Service) idzie o krok dalej i polega na udostępnianiu gotowych aplikacji.

DaaS (Data as a Service) jest trochę inną usługą polegającą na umożliwieniu bardziej bezpośredniego pobierania danych z chmury.

Dostawcami usług Big Data w chmurze zostawali najczęściej albo organizacje korzystające z Big Data do własnych celów (Google, Amazon) albo dostawcy rozwiązań bazodanowych (Microsoft, Oracle). [16-18] Poniżej znajduje się lista wybranych rozwiązań Big Data w chmurze [13-16]:

- Amazon EC2
- Google Cloud Platform
- IBM Big Data Cloud
- Microsoft Azure
- Oracle Cloud

Z perspektywy użytkownika końcowego w tym specjalisty ochrony zdrowia. Najprostszymi usługami będą SaaS, IaaS (np. Amazon EC2) oraz DaaS, natomiast PaaS będzie służyło zespołom, które cenią sobie programowalność środowiska (np. Windows Azure). (rycina 2.)



Rycina 2. W ramach *cloud computing* oferowane są następujące modele eksploatacji [16-18]

PIŚMIENNICTWO

- Schmarzo B. Big Data. Understandinga How Data Powers Big Business. Indianapolis; John Wiley & Sons Inc., 2013.

- Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. *Health Inform Sci Syst* 2014; 2: 3-10.
- Kryśkiewicz Ł.: Jak wygląda system e-usług w Danii. [cytowany 27 sierpnia 2016]. Adres: <http://di.com.pl/jak-wyglada-system-e-uslug-w-danii-54770>
- Papińska-Kacperek J. Usługi cyfrowe. Perspektywy wdrożenia i akceptacji cyfrowych usług administracji publicznej w Polsce. Łódź; Wyd. Uniwersytetu Łódzkiego, 2013.
- Diebold F. On the Origin(s) and Development of the Term "Big Data". Pier working paper archive. Pennsylvania; Penn Institute for Economic Research, Department of Economics. University of Pennsylvania, 2012.
- Mach-Król M. Analiza i strategia Big Data w organizacjach. *Stud Mater Pol Stow Zarz Wiedzą* 2015; 74: 43–55.
- Bollier D. The promise and Peril of Big Data. Raport of the Aspen Institute. Washington; Communications and Society Program, 2010.
- Marconi K, Dobra M, Thompson C. The use of Big Data in Healthcare. In: Liebowitz J. Big Data and Business Analytics. Boca Raton; CRC Press, 2013: 229–248.
- Groves P, Kayyali B, Knott D, Kuiken S. The big data revolution in healthcare. London; McKinsey&Company, 2013
- Mateos A, Rosenberg J. Chmura obliczeniowa Rozwiązania dla biznesu. Gliwice; Wyd. Helion, 2011.
- Reczek E. Zastosowanie chmury obliczeniowej w sektorze ochrony zdrowia. *Zesz Nauk WSZIB* 2014;33: 1-11.
- AlZain MA, Pardede E, Soh B, Thom JA. Cloud Computing security: From single to multi-clouds. *Proceedings of the HICSS*, 2011.
- Fiedler A, Brown I. Chmury obliczeniowe – ekspertyza. Bruksela; Dyrekcja Generalna ds. Polityki Wewnętrznej Unii Europejskiej, 2012.
- Kroes N. Setting up the European Cloud Partnership. Speech at the World Economic Forum, Davos 2012.
- Burnejko M. Kto korzysta z chmury w Polsce (i na świecie). Blog ekspertów IT. 04 września 2015. [cytowany 8 sierpnia 2016]. Chmurowisko.pl, <http://chmurowisko.pl/>
- Mapping the cloud maturity curve: The Economist Intelligence Unit. 2015. 18 luty 2016. [cytowany 8 sierpnia 2016]. http://www.corporate-leaders.com/sitescene/custom/userfiles/file/White_Papers/Mapping%20the%20cloud%20maturity%20curve.pdf, 25.08.2016. 34.
- Mell P, Grance T. The NIST Definition of Cloud Computing. 2011. 18 luty 2016. [cytowany 8 sierpnia 2016]. <http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication80-145.pdf>,
- Uptake of Cloud in Europe. Raport IDC na zlecenie Komisji Europejskiej, 2013. 18 luty 2016. [cytowany 8 sierpnia 2016]. <https://ec.europa.eu/digital-single-market/en/news/final-report-study-smart-20130043uptake-cloud-europe>.